**Návrh národní koncepce dlouhodobé ochrany digitálních dat pro knihovny**

6.11.2014

Obsah

[1 Uvedení do problému 3](#_Toc403076971)

[1.1 Kulturní digitální data v knihovnách 3](#_Toc403076972)

[1.2 Současná legislativa 3](#_Toc403076973)

[1.3 Dlouhodobá ochrana digitálních informací jako oblast výzkumu a praxe 4](#_Toc403076974)

[1.4 Současný stav v ČR 5](#_Toc403076975)

[1.5 NDK standard 5](#_Toc403076976)

[1.6 Národní koordinace LTP v jiných zemích 6](#_Toc403076977)

[1.7 Česká situace jako východisko pro formulaci strategie 7](#_Toc403076978)

[2 Definice cílů a principů národní strategie 9](#_Toc403076979)

[3 Ideální stav za pět let v oblasti dlouhodobé ochrany 11](#_Toc403076980)

[3.1 Jádro systému pro dlouhodobou ochranu knihovních dat 11](#_Toc403076981)

[3.2 Další certifikované systémy 12](#_Toc403076982)

[3.3 Otevřená řešení 12](#_Toc403076983)

[3.4 Born-digital data 12](#_Toc403076984)

[3.5 Vzdělávání, výzkum a certifikace 13](#_Toc403076985)

[3.6 Sdílení a financování 13](#_Toc403076986)

[4 Konkrétní doporučená opatření 14](#_Toc403076987)

[4.1 Plán na zajištění dlouhodobé ochrany digitálního kulturního a vědeckého dědictví v ČR 14](#_Toc403076988)

[4.2 Státem garantovaná datová centra pro zajištění bitové ochrany 14](#_Toc403076989)

[4.3 Státem garantovaný systém pro zajištění logické ochrany 15](#_Toc403076990)

[4.4 Postupné budování sítě certifikovaných úložišť v jednotlivých institucích 15](#_Toc403076991)

[4.5 Podpora Open Source řešení pro zajištění logické ochrany 15](#_Toc403076992)

[4.6 Zřízení meziresortního metodického centra pro dlouhodobou ochranu 16](#_Toc403076993)

[4.7 Krajské knihovny jako producenti a správci digitálních dat 16](#_Toc403076994)

[4.9 Finanční mechanismy 17](#_Toc403076995)

[4.10 Organizační a systémové změny 17](#_Toc403076996)

[5 Doporučení v oblasti zajištění kvality, standardizace a certifikace 19](#_Toc403076997)

[5.1 Zajištění kvality; audit a certifikace 19](#_Toc403076998)

[5.2 Doporučené postupy 20](#_Toc403076999)

[6 Postup realizace koncepce 20](#_Toc403077000)

[7. Současné aktivity 21](#_Toc403077001)

[8. Použité zdroje 22](#_Toc403077002)

# 1 Uvedení do problému

## 1.1 Kulturní digitální data v knihovnách

Knihovny se musí dlouhodobou ochranou (někdy používán termín dlouhodobá archivace) digitálních informací zabývat především proto, aby v digitální době přežily jako klíčové informační instituce. Jejich klienti dnes přicházejí především po síti a očekávají plnohodnotné služby v digitální a online podobě. Pokud si knihovny chtějí udržet svou tradiční roli, musí se přizpůsobit a poskytovat online jak aktuální informační zdroje (článkové databáze, webové archivy), tak i dokumenty, které vznikají výhradně v digitální podobě (born-digital) a nahrazují tak dokumenty klasické. I tyto jsou součástí kulturního dědictví. Knihovny musí born-digital dokumenty vybírat, ochraňovat a zpřístupňovat. Vedle toho uživatelé automaticky předpokládají, že knihovny online (tedy v digitalizované podobě) nabídnou i dokumenty, které jsou součástí jejich sbírek již z minulosti. Vedle výše uvedeného knihovny, např. univerzitní, musí stále častěji přijímat a tedy zpřístupňovat i nové typy objektů (vědecká data, audiovizuální dokumenty) nebo poskytovat nové služby (publikační, archivace, nové metody akvizice).

Knihovny byly po staletí zvyklé pracovat s fyzickými dokumenty. V digitálním světě takto dlouhou zkušenost zatím nemají. Pokud ale chtějí knihovny jako instituce přežít, musí se digitálnímu světu přizpůsobit. Motivace zabývat se dlouhodobou ochranou přichází také zvenčí. Česká republika by jako členská země EU měla následovat evropská doporučení. Dlouhodobá archivace kulturního dědictví v digitální podobě je jednou z klíčových aktivit Digitální agendy pro Evropu[[1]](#footnote-1), která pokrývá jak oblast digitalizace, tak dlouhodobou ochranu a zpřístupnění.

## 1.2 Současná legislativa

Povinnost ukládat a ochraňovat dokumenty uložené jako sbírky pro budoucí generace v prostředí knihoven pramení z knihovního zákona (Zákon o knihovnách a podmínkách provozování veřejných knihovnických a informačních služeb č. 257/2001 Sb., ve znění pozdějších změn). Tento zákon se vztahuje na knihovny, které jsou zapsány v evidenci knihoven MK ČR. V paragrafu 18 se mluví o ochraně knihovního fondu. Knihovní fond a knihovní dokumenty musejí být ochráněny před poškozením a zcizením; uloženy ve vhodných podmínkách a provozovatel knihovny musí zajistit převod dokumentu na jiný druh nosiče, pokud je to potřeba k jejich uchování.

Knihovní zákon explicitně nemluví o knihovních fondech a dokumentech, které mají digitální formu. Z definice pojmu Knihovní fond a Knihovní dokument v § 2. zákona lze ale usuzovat, že tyto digitální formu zahrnují[[2]](#footnote-2). Knihovní sbírky obsahují digitální dokumenty dnes běžně.

Podobná situace panuje i v dalších kulturních paměťových institucích. V oblasti archivů např. Archivní zákon (Zákon o archivnictví a spisové službě a o změně některých zákonů č. 499/2004 ve znění novely 167/2012) vymezuje v paragrafu 2 bodě b) pojem „péče o archiválie“, který zahrnuje vedle výběru, evidence, zpracování také ochranu archiválií. Archiválie je dle bodu f) stejného paragrafu „takový dokument, který byl vzhledem k době vzniku, obsahu, původu, vnějším znakům a trvalé hodnotě dané politickým, hospodářským, právním, historickým, kulturním, vědeckým nebo informačním významem vybrán ve veřejném zájmu k trvalému uchování a byl vzat do evidence archiválií, …“.

Z obou zákonů tedy vyplývá, že se knihovny a archivy musejí zabývat ochranou dokumentů, které existují a do institucí přicházejí pouze v digitální formě.

## 1.3 Dlouhodobá ochrana digitálních informací jako oblast výzkumu a praxe

Pro trvalé uchování digitálního kulturního dědictví je třeba implementovat organizační procesy a technické informační systémy, které umožní vybudování dlouhodobých důvěryhodných digitálních úložišť, tj. systémů, které jsou schopny zajistit udržení a použitelnost digitálních informací v horizontu stejném, na jaký jsme zvyklí u fyzických dokumentů – po dobu několika set let.

Dlouhodobá ochrana digitálních dat se jako problém objevila už v polovině 80. let 20. století, kdy si někteří archiváři začali uvědomovat, že s udržením digitálních dat to nebude tak jednoduché: že ochranu digitálních dat je třeba dobře plánovat a je třeba počítat s neustálými zásahy a přesuny dat, protože digitální technologie mají krátkou životnost, např. [FOX, 1985]. Velmi důležitou byla americká zpráva *Preserving Digital Information*, která vznikala od roku 1994 a byla vydána v roce 1996 [GARRETT a WATERS, 1996]. Jako jeden z prvních dokumentů definovala dlouhodobou ochranu digitálních dat v dnešním smyslu, popsala problematiku zastarávání SW i HW, možnosti řešení (migrace), otázky integrity digitálních dokumentů, jejich provenience a zachycení kontextu atp. V závěru zprávy autoři konstatovali, že v době jejího vydání neexistovala žádná vhodná metoda na dlouhodobou ochranu a společnost a její instituce nejsou na tuto problematiku připraveny. V archivářské komunitě na sebe nedala reakce dlouho čekat, a v roce 1999 vznikl první draft referenčního rámce OAIS (Open Archival Information System), pozdější normy ISO 14721:2002 resp. 2012. Ta je dodnes základním orientačním bodem pro každého, kdo se dlouhodobou ochranou seriózně zabývá. Další vývoj pak vedl ke standardizaci konceptu dlouhodobého důvěryhodného digitálního archivu (viz ISO 16363:2012 a prováděcí norma ISO 16919).

Dlouhodobá ochrana digitálních informací (anglicky nejčastěji *digital preservation* nebo v širším kontextu správy digitálních dat označovaná jako *digital curation*) je výzkumnou oblastí[[3]](#footnote-3) a v posledních letech také oblastí každodenní praxe v řadě institucí.

V digitálním světě, stejně jako ve světě fyzických dokumentů, není možné uchovávat všechny dokumenty. Knihovny, i v rámci např. povinného výtisku, musí jasně definovat, které objekty mají být předmětem dlouhodobé ochrany a musí rozlišovat, která data skutečně má smysl ukládat dlouhodobě a s vynaložením maximální péče (born-digital, digitized masters) a která ukládat dočasně (uživatelské kopie).[[4]](#footnote-4)

Nutným předpokladem a prvním krokem k dlouhodobé ochraně digitálních dokumentů a vytvoření dlouhodobého důvěryhodného digitálních archivu je zajištění ochrany bitů. Knihovny, které mají vlastní datová úložiště a datová centra vědí, že to znamená zajištění neměnnosti dat, fyzickou ochranu, zálohování a vícenásobné ukládání na více médiích ve více lokalitách.

Druhým krokem je zajištění logické ochrany, jejímž cílem je zajištění použitelnosti, srozumitelnosti, zobrazitelnosti archivovaných digitálních dokumentů v jakémkoliv okamžiku v budoucnu i navzdory změnám technologií. Toho je dosaženo vytvářením a ukládáním technických, ochranných, administrativních metadat po celou dobu existence dokumentu. Tato metadata jej popisují, udržují kontext a lze je využít na ochranné aktivity jako např. migrace formátu apod. Logická ochrana předpokládá nejen technické systémy, ale hlavně nastavení procesů podle požadavků výše uvedených norem.[[5]](#footnote-5)

Vedle standardizace organizačních procesů je také nutné mít zajištěné financování, odborné pracovníky, odpovídající prostory aj. To jsou vše příklady požadavků z normy ISO 16363 na to, aby mohlo být digitální úložiště certifikováno jako důvěryhodné.

## 1.4 Současný stav v ČR

V České republice se knihovny dlouhou dobu zaměřovaly primárně na produkci nebo sklízení digitálního obsahu - projekty financované z veřejných rozpočtů podporovaly generování digitálních dat (digitalizace, archivace webu), přičemž dlouhodobá ochrana těchto data byla podceňována. V současnosti je v této oblasti situace výrazně fragmentovaná, projekty mají různé cíle, nejčastěji zpřístupnit data v digitalizované podobě, vzácně ochránit digitální data pro budoucnost. Chybí jednotná vize Ministerstva kultury (dále jen MK ČR) a následně samotných knihoven pro oblast dlouhodobé ochrany. Z velkých digitalizačních projektů probíhajících v ČR (NDK[[6]](#footnote-6), Google Books[[7]](#footnote-7), Manuscriptorium[[8]](#footnote-8)) řeší dlouhodobou ochranu explicitně pouze NDK. Dále existují menší digitalizační aktivity (VISK, Krajské digitalizace, Norské fondy), které o dlouhodobé ochraně vyprodukovaných dat alespoň uvažují. Aktivní na tomto poli je vedle Národní knihovny ČR (dále jen NK ČR) a MZK také Knihovna Akademie věd.

Velkým úspěchem je prosazení standardní podoby výstupu digitalizace (tzv. NDK standard) na národní úrovni. Data a metadata vytvořená dle tohoto standardu jsou do jisté míry na dlouhodobou ochranu připravena, i když nebudou uložena v odpovídajícím long-term preservation (dále LTP) systému.[[9]](#footnote-9)

Množství projektů, ve kterých digitální data vznikají, vede k tomu, že různé projekty aplikují různé metody zajištění kvality digitalizace a ukládání. Hlavním problémem je neexistence explicitního požadavku zajistit ve všech projektech odpovídající trvalé uložení klíčových digitálních dat. Chybějí jasné definice dat, které je nutné dlouhodobě ochránit (z hlediska kvality reprodukce například), a také jasná definice způsobu této ochrany, tj. požadavků na systémy, ve kterých mají být tato data ukládána.

Všechna digitální data produkovaná a shromažďovaná v projektech financovaných z veřejných rozpočtů by měla podléhat explicitním požadavkům na zajištění kvality dlouhodobé ochrany, tj. měla by být definována data, která tvoří jádro kulturního dědictví (např. born-digital data; data z archivace webu; určité typy digitalizovaných dat) a která je třeba uchovávat v certifikovaných dlouhodobých digitálních repozitářích. V současné době má každý projekt vlastní představy o kvalitě, způsobu uložení dat a metadat.

Neznamená to, že bychom volali po jednotném systému a jednom úložišti, naopak diverzita přístupů a systémů sníží závislost a riziko. Jde spíš o koordinaci na úrovni zajištění kvality (certifikace úložiště, standardy pro data, procesy), financování a metodik, tedy jasný návod pro rozlišení těch dat, kterým je třeba věnovat maximální péči.

## 1.5 NDK standard

Prosazení výše zmíněného NDK standardu jako národního standardu pro digitalizaci v knihovnách považujeme za silnou stránku současného stavu. Formulace standardu NDK navazovala na zkušenosti s předešlou generací národního standardu postaveného na starších formátech DTD. Archivní balíčky v DTD standardu byly dlouhá léta generovány v digitalizačním procesu novodobých publikací v Národní knihovně a v řadě dalších institucí. Standard postavený na DTD pro periodika a pro monografie byl závazný pro digitalizaci v rámci projektů VISK7, následovala ho ale mj. také Knihovna akademie věd a další. Firmy jako Elsyst upravovaly své nástroje tak, aby byly schopné metadata a data v určené podobě vytvářet. Data v těchto strukturách jsou často velmi cenná, jde především o skeny nejohroženejších periodik.

Standard DTD byl jak v NK, tak v rámci VISK7 nahrazen NDK standardem v roce 2013[[10]](#footnote-10) (v letech 2011 a 2012 bylo v rámci VISK povoleno používat oba standardy). Současný standard, balíček NDK, je založen na v současnosti obvyklých a doporučovaných formátech metadat (METS, PREMIS, MODS, Dublin Core, MIX a ALTO XML). Všechny jmenované standardy jsou de-facto standardy, vyvíjené a podporované Kongresovou knihovnou USA. Jsou bez výjimky využívány ve všech projektech dlouhodobé archivace v knihovnách. Metadatová specifikace vznikala na základě podobných projektů a ve spolupráci s národními knihovnami Nizozemí, Norska, Finska a Kongresové knihovny, USA.

Jejich prosazení na národní úrovni a uplatnění v projektech VISK, Krajské digitalizace a jinde považujeme za velký úspěch. Ať se jedná o jakýkoliv projekt, základní úroveň ochrany je použitím tohoto standardu zajištěna, stejně jako tolik žádaná konsistence zajišťující kooperabilitu mezi projekty a jejich repozitáři. NDK standard nejsou pouze metadata, stejná pozornost byla věnována specifikacím obrazových dat a vytváření profilu JPEG 2000 a OCR souborů v METS ALTO. Standard NDK odpovídá současné nejlepší praxi v paměťových institucích ve světě. Ukládání jednotlivých naskenovaných stran jako digitalizačních masterů ve formátu JPEG 2000 představuje posun k formátu, který poskytuje lepší vizuální možnosti reprezentace obsahu než starší JPEG nebo TIFF. Je také prokazatelně odolnější vůči bitovým ztrátám než jiné formáty (PNG, JPEG) a výrazně (tj. několikanásobně) šetří prostor oproti formátu TIFF při zachování kvality. Ačkoli nejde o formát masově rozšířený, jeho obliba jako archivního formátu pro digitalizovaná data roste a je používán v řadě knihoven a archivů (Britská knihovna, Americká kongresová knihovna, Britský národní archiv, Národní knihovny Nizozemí, Norska, Wellcome Library, a mnoho dalších). Mnohé instituce provádějí převod z formátu TIFF na JPEG 2000 u svých archivních dat, často z důvodu úspory prostoru na datovém úložišti a také z důvodu větší flexibilnosti formátu JPEG 2000. TIFF i JPEG 2000 jsou považovány za rovnocenné z pohledu vhodnosti pro dlouhodobou ochranu.

Samozřejmě, národní standard NDK je standardem pro digitalizovaná data, a pro archivaci e-born dokumentů (deposit elektronických publikací, E-knih, šedé literatury, web archivu atd.) je třeba pracovat na dalším rozvoji tohoto standardu. Jako základ by mě být i pro archivní účely využíván METS jako tzv wrapper, který obsahuje metadata ve schématech PREMIS a MODS, DC. Samozřejmě s tím, že archivní balíček oproti balíčku z digitalizace, by měl být obohacen o technická a administrativní metadata potřebná pro efektivní ochranu daného typu materiálu.

## 1.6 Národní koordinace LTP v jiných zemích

V různých zemích ke koordinaci aktivit spojených s dlouhodobou archivací na národní úrovni přistupují velmi odlišně. EU v rámci svých aktivit, které jinak doposud směřovaly spíše k podpoře výzkumu a vývoje v oblasti dlouhodobé archivace, realizovala šetření v rámci projektu ENUMERATE, kde byly mj. pokládány otázky týkající se dlouhodobé archivace. V podstatě to jsou jedna z mála tzv. tvrdých dat o praxi v jednotlivých institucích v EU a ukazují, že evropské paměťové instituce sice aktivně skenují a pořizují digitální data, ovšem dlouhodobou archivací se skutečně zabývá nejvýše jedna čtvrtina z nich[[11]](#footnote-11).

V různých zemích probíhá koordinace dlouhodobé archivace na národní úrovni různě. Někde se snaží koordinovat aktivity v archivech i knihovnách (Kanada, Nový Zéland a částečně také Holandsko), obvykle v rámci celonárodního projektu přechodu na digitální agendu v celém státě. Jinde knihovny více spolupracují s akademickou sférou, a společné projekty vedou buď ke sdílení znalostí a podpoře výzkumu (Velká Británie a organizace jako Digital Preservation Coalition (DPC); Holandsko, Polsko). Někde je pro dlouhodobou archivaci v paměťových institucí klíčová aktivity některé dominantní instituce (Dánsko, Norsko, Francie, Finsko). V některých zemích je zájem o dlouhodobou archivaci diktován požadavky a projekty shora (Slovenský projektu OPIS a v rámci něj projekt CDA), jiné země nebo instituce budují systémy zdola (Sasko a Německo obecně, Polsko). Příkladem inspirativním pro českou republiku může být Polsko nebo některé německé země, kde využívají kapacity a dovednosti akademické sféry a akademická výpočetní centra k budování služeb pro paměťové instituce.

Zvláštní postavení mají Německo, Británie a Spojené státy. V těchto zemích se dlouhodobé ochraně digitálních informací věnuje více než jedna instituce a znalostních center je tak více. V ostatních zemích jde většinou o národní instituci typu knihovna nebo archiv, případně nějaká jiná centrální instituce. V Německu, USA a Velké Británii jsou vedoucími instituce jak národní knihovny, tak i národní archivy. Celou síť dále doplňují aktivní univerzity a datová centra (např. pro vědecké informace). Lze říci, že koordinace v těchto zemích neprobíhá.

## 1.7 Česká situace jako východisko pro formulaci strategie

* Díky existenci programu VISK má v České republice řada institucí zkušenosti s digitalizací (byť často externí, dodavatelskou) novodobých i historických fondů. Také další aktivity a finanční zdroje byly v České republice věnovány obvykle na pořizování digitálních dat (Norské fondy, atd.) nebo jejich zpřístupňování (Digitální knihovna Kramerius). Tyto finanční zdroje vždy vyžadují po institucích dodržení standardů pro tvorbu metadat i formáty digitalizace. Projekt digitálního úložiště pro muzea a galerie CITEM byl zcela zapomenut.
* V České republice existuje silná infrastruktura akademické sítě a jsou budována akademická úložiště (CESNET). Akademická sféra se věnuje především zpřístupňování výsledků vědecké činnosti a kvalifikačních prací s využitím open source technologií (repozitáře DSpace, Invenio, apod.). Bohužel například aktivity spojené s archivací vědeckých dat (především tzv. malých vědeckých dat z oborů *Arts, humanities and social sciences*), obdobné v zahraničí, jsou u nás nepříliš časté. Vybudování datového archivu podobného např. britskému UK Data Archive, holandskému DANS nebo německému GESIS (Lipsko) není ze strany relevantního ministerstva nijak podporováno. Spolupráce mezi paměťovými institucemi (knihovny, muzea a galerie, a archivy) a akademickou sférou v oblasti dlouhodobé archivace je stále v plenkách. Těžko najdeme společné projekty výpočetního centra nějaké univerzity, a krajského muzea.
* Česká republika má jednu nejstarších tradic archivace internetu v Evropě. Objemy dat z archivace webu jsou enormní a financování ochrany těchto dat by mělo dostat jasný právní rámec (statut WebArchivu, jasné přidělení odpovědnosti, závazek státu financovat ochranu těchto dat) s tím, že WebArchiv by měl směřovat k prokázání kvality své činnosti (dodržení ISO normy pro kvality sbírek webového archivu – ISO 14 873, a dodržení ISO 28500).
* České kulturní instituce mají ve svých sbírkách data vznikající od roku 1996 (digitalizace), případně data z archivace webu od roku 2001. Tato data jsou z technického hlediska velmi stará, dlouhodobá ochrana digitálních informací přesto donedávna neměla pevné místo v žádné z českých institucí.
* Vedle menších projektů financovaných z rozpočtů Ministerstva kultury v rámci programu VISK je zde rozhodující role Národní knihovny a MZK a jejich projektu NDK, který usiluje u vybudování dlouhodobého důvěryhodného úložiště v souladu s ISO 16363.
* Role Ministerstva kultury je pro paměťové instituce a jejich snahy zabývat se dlouhodobou archivací zcela klíčová – jsou často jejich zřizovateli, nebo vytváří finanční nástroje a pravidla, které by mohly aktivitu v oblasti stimulovat. Financování (VISK apod.) šlo doposud vždy společně s tlakem na dodržení standardů a do budoucna by mělo být provázeno požadavkem zajištění dlouhodobé archivace.
* Role Ministerstva vnitra je specifická a dlouhodobá archivace v oblasti archivnictví je dána zákonem. Ovšem i zde existují oblasti, kde se skenují a zpřístupňují historické nebo speciální fondy, bez toho, že by jim bylo zajištěna dostatečná úroveň dlouhodobé ochrany – a to často i v masové míře – jako v Ústavu pro studium totalitních režimů.
* Ohnisek zájmu o dlouhodobou ochranu digitálních informací je nyní více (KNAV, NFA, NA, MZK, CESNET, MUNI, FAMU, CITEM, NDK) a tento zájem je třeba kultivovat a podpořit. Díky neexistenci podpory z ministerstev či vlády, jde o zájem zdola. Také díky tomu se důvody tohoto zájmu, někdy i podstatně, liší.
* Konkrétní instituce mají skrz některé své zaměstnance velmi dobré znalosti problematiky dlouhodobé ochrany digitálních informací, kooperace se zahraničím začala již po roce 2000 a stále pokračuje.
* Z výše uvedeného by mělo být zřejmé, že dlouhodobá archivace digitálních dokumentů je oblast, která spadá do gesce institucí financovaných z rozpočetů minimálně tří ministerstev: Ministerstva kultury, Ministerstva vnitra a Ministerstva školství, mládeže a tělovýchovy.

# 2 Definice cílů a principů národní strategie

Základním cílem pro příštích pět let by mělo být vytvoření podmínek pro snížení bariér a diverzifikaci přístupu k dlouhodobé ochraně při jasném stanovení standardů pro tuto oblast.

K tomuto cíli bychom se měli dostat postupným rozvíjením aktivit v jednotlivých oblastech:

* **Rozvoj explicitního a měřitelného plánování** - na úrovni institucí; strategické plánování na úrovni systému knihoven a jednotlivých projektů.
* **Zlepšení organizačních podmínek** - pro dlouhodobou ochranu v knihovnách a to jak na úrovní jednotlivých knihoven (standardizace procesů a správy dat, plánování a dokumentace vyžaduje další oddělení, pracovní síly, finance), tak na úrovni systému knihoven (například vznikem koordinačního centra viz níže).
* **Podpora spolupráce a sdílení -** jak na domácí úrovni mezi paměťovými institucemi různých resortů, s univerzitami a výzkumnými organizacemi atd., tak na mezinárodní úrovni s jinými paměťovými institucemi. Sdílení procesů, plánů, metodik, HW i SW infrastruktury.
* **Snížení bariér** - zajistit, aby každá instituce, která produkuje nebo sbírá digitální data, měla přístup k znalostem, službám nebo konkrétním technologiím. Manažer každé instituce musí mít takové znalosti nebo dokumentaci, aby mohl rozhodnout, jak se bude dlouhodobou ochranou zabývat, jak budou data ukládána, a měl by mít možnost využít služeb systémů jiných institucí, nebo státem poskytovaných služeb.
* **Rozvoj mechanismů financování dlouhodobé ochrany** - vytvoření udržitelného a kontrolovatelného finančního nástroje pro financování projektů v oblasti dlouhodobé ochrany, včetně jednorázových akcí. Musí existovat závislost a návaznost projektů digitalizace na toto financování. Mechanismus by měl rozlišit prostý nákup HW, implementaci systémů pro správu dat, jednorázové akce ochrany (migrace z HW na HW, formátové migrace), komplexní logickou dlouhodobou ochranu a výzkumné projekty.
* **Podpora vzdělávání, kvalifikace a výzkum** - existující formální univerzitní obory typu „Knihovnictví a informační věda“, “Archivnictví” a “Muzeologie” a některé v oblasti informačních technologií by měly integrovat problematiku dlouhodobé ochrany digitálních dat do svých učebních programů, případně vytvořit nová zaměření na *digital curation* a dlouhodobou ochranu např. podle doporučení projektu DigCurV[[12]](#footnote-12) nebo podle již existujících učebních programů v zahraničí. Je také žádoucí specifická podpora projektů spolupráce a výzkumu v oblasti dlouhodobé ochrany (knihovny a univerzity), systematická podpora studentských stáží a akademických prací.
* **Centrální metodické centrum a zavedení standardů a metodik** - pro zajištění kvality v oblasti dlouhodobé ochrany a zavedení národního mechanismu externí certifikace důvěryhodného dlouhodobého úložiště pro data v knihovnách. To musí být podpořeno fungujícím systémem a metodikou tzv. sebehodnocení, např. dle DSA (Data Seal of Approval[[13]](#footnote-13)) nebo Nestor Seal[[14]](#footnote-14). Vytvoření doporučení pro instituce, jak má vypadat institucionální/sbírková strategie dlouhodobé ochrany, jak postupovat ve snaze dosáhnout statutu důvěryhodného dlouhodobého repozitáře. Tento proces musí být iniciován a udržován z pozice MK ČR, ideálně skrz nově zřízené *centrální metodické centrum pro dlouhodobou ochranu*. Toto metodické centrum může vzniknout při již existující instituci, nebo může být vybudováno zcela samostatně. Institut by měl být financovaný nezávisle na případné hostitelské instituci, a to přímo z prostředků MK ČR.
* **Důraz na advokacii dlouhodobé ochrany** - pro úspěch dlouhodobé ochrany jsou podle nás klíčové jak finanční prostředky, tak odhodlání lidí se touto problematikou zabývat. Protože dlouhodobá ochrana je finančně náročná činnost probíhající někde v pozadí, často daleko od každodenního provozu digitální knihovny, je třeba tyto činnosti popularizovat a vysvětlovat jejich význam pro zachování služeb knihoven.

# 3 Ideální stav za pět let v oblasti dlouhodobé ochrany

Existuje systém pro dlouhodobou ochranu knihovních dat. Nejedná se o systém v technickém slova smyslu, ale o soubor pravidel, institucí, technických zařízení a aplikací, vzdělávacích programů a organizační rámec okolo nich. Systém může být součástí širšího konceptu pro ochranu digitálních informací v ČR, nejen knihovnických ale např. i dat z archivů, muzeí, galerií, vědeckých institucí apod. Jeho hlavními komponenty jsou:

* Hardwarová a softwarová řešení
	+ státem garantované datové centrum pro zajištění bitové ochrany
	+ státem garantovaný LTP systém pro zajištění logické ochrany digitálních dat
	+ lokální řešení (Open Source nebo komerční) jednotlivých institucí různé úrovně a kvality
	+ DÚ CESNET, výpočetní centra univerzit
* Instituce
	+ mezi resortní metodické centrum pro LTP
	+ centrální datové centrum jako samostatný subjekt
	+ jednotlivé knihovny, univerzity, archivy, muzea
* Finanční mechanismy
	+ stávající mechanismy jako VISK obohacené o nové prvky
	+ nové finanční mechanismy financující LTP aktivity, vlastní aktivity institucí nebo využití centrální infrastruktury /CESNET/centrální datové centrum
* Standardy
	+ standardy dat a metadat
	+ doporučení pro LTP, LTP metodiky
	+ ISO normy

## 3.1 Jádro systému pro dlouhodobou ochranu knihovních dat

Jádro systému pro dlouhodobou ochranu knihovních dat tvoří státem garantovaná data centra poskytovaná např. CESNETem a MK ČR (případně jiné ministerstvo pokud půjde o ještě širší řešení – Ministerstvo vnitra ČR, které má ve své gesci archivy atd.)

* + Data centra mají více technologií ukládání a více modelů využití.
	+ Infrastruktura umožňuje distribuci dat mezi oblastmi, jsou dostupná rychlým datovým připojením (ne převážení pasek); jsou flexibilní a instituce nakupují kapacity podle potřeb.
	+ Existuje dohoda o automatickém převzetí odpovědnosti za data vznikající za veřejné finance - pokud některá instituce nechce dlouhodobou ochranu dat řešit a vytváří digitální data, musí mít plán na jejich ochranu. Tedy musí vědět komu, jak a v jaké podobě má data předat. Podobně musí být plánováno pro případ zániku, instituce, musí být jasné, kdo je za uložená data odpovědný.
	+ Pro data kterékoli knihovny je zajištěna minimálně ochrana bitů - každá instituce má možnost použít CESNET nebo datové centrum MK ČR pro zajištění bitové ochrany (národní cloud v podobě IaaS – Infrastructure as a Service, resp. SaaS – Storage as a service).

## 3.2 Další certifikované systémy

Některé z velkých institucí nebo konsorcií (NK ČR, MZK, kraje, univerzity jako např. Masarykova univerzita, Karlova univerzita, CESNET, nově zřízená instituce jako např. Datový archiv) mají k dispozici certifikovaná řešení pro dlouhodobou ochranu digitálních dat včetně adekvátního financování a personálu.

* + Existující Long-term Preservation (LTP) systémy mají povinnost převzetí dat od menších institucí v případě zániku a smlouvy mezi sebou o tomtéž.
	+ Standardizace formátů a metadat napříč institucemi stejného typu (i různé velikosti).
	+ Standardizace pracovních postupů a dokumentace.
	+ Některé instituce používají externí cloudové úložiště nebo služby cloudového LTP systému[[15]](#footnote-15).
	+ Některé z institucí poskytují přímo služby dlouhodobé ochrany jiným institucím (DPaaS – digital preservation as a service).

## 3.3 Otevřená řešení

Vedle vždy existující možnosti pořízení komerčního LTP systému, existuje státem garantované a lehce dostupné open source řešení systému na dlouhodobou ochranu pro malé a střední instituce. Instituce mají na výběr - buď využijí existující open source řešení nebo mohou využít službu některého z provozovaných systémů ve větších institucích. Minimálně ochrana na úrovni bitů je garantována státem pro všechna data vznikající v projektech financovaných z veřejných zdrojů.

**Obr. 1. Komponenty národního systému pro ochranu digitálních dat**

## 3.4 Born-digital data

* + NK ČR nebo jiné pověřené knihovny mají ověřený mechanismus na sběr/shromažďování/vybírání born-digital dokumentů jako povinného výtisku a zajišťují jejich logickou dlouhodobou ochranu.
	+ NK ČR nebo jiné pověřené knihovny pravidelně sbírají obsah českého Internetu (doména .cz a další weby bohemikální povahy) a zajišťují logickou dlouhodobou ochranu těchto dat. WebArchiv prochází pravidelně auditem kvality podle příslušné ISO normy.

## 3.5 Vzdělávání, výzkum a certifikace

* + Součástí infrastruktury je centrální metodické centrum pro dlouhodobou ochranu, které poskytuje vzdělávání, poradenství, standardy, metodiky a certifikační služby včetně např. školení auditorů.
	+ Univerzitní vzdělávací programy jsou adekvátně obohaceny o problematiku dlouhodobé ochrany digitálních dat.
	+ Dlouhodobá ochrana digitálních dat je stálou, reálně prováděnou a metodicky a finančně podporovanou agendou MK ČR, a dalších ministerstev, v souladu s doporučeními EU, a to nejen na papíře.
	+ V definicích pracovních pozic ve veřejných institucích jsou jasně popsány pracovní pozice a kvalikační požadavky na kurátory digitálních dat, digitální knihovníky, správce digitální knihovny, metadatové specialisty a analytiky pro dlouhodobou archivaci.

## 3.6 Sdílení a financování

* + Ministerstvo kultury nabízí dlouhodobě udržitelné mechanismy financování dostupné pro projektové aktivity, vzdělávání v oblasti dlouhodobé ochrany a LTP systémů; experimenty a hodnocení vznikajících LTP systémů, nástrojů atd. v koordinaci s MV ČR a MŠMT.
	+ Knihovny programově spolupracují s dalšími paměťovými institucemi, jako jsou archivy, muzea, galerie apod., se kterými mohou sdílet HW/SW infrastrukturu a další řešení. Podobná spolupráce existuje také s univerzitami. Je podpořena přílišnými finančními mechanismy.

# 4 Konkrétní doporučená opatření

## 4.1 Plán na zajištění dlouhodobé ochrany digitálního kulturního a vědeckého dědictví v ČR

* Z pozice vlády ČR prostřednictvím relevantních ministerstev (MK, MV, MŠMT) vytvořit plán na zajištění dlouhodobé ochrany digitálního kulturního a vědeckého dědictví v ČR.
* Definovat cíle, organizační opatření, pravidla, rozhodnutí přidělení kompetencí (ochrana bit streamu, logická ochrana) pro oblast knihoven, archivů, muzeí, galerií ale také univerzit a vědecká data.
* Společná strategie MK, MV a MŠMT by měla:
	+ Definovat co jsou digitální kulturní a vědecká data trvalé hodnoty.
	+ Definovat co je dlouhodobá ochrana digitálních dat a jaké procesy budou využity k jejímu zajištění.
	+ Odvolat se na české verze norem ISO 16363, ISO 14721, a normu ISO 14873, zmíněny by měly být i normy ISO 27000 a ISO 9000 a další standardy (DSA, Nestor Checklist) pro začínající instituce.
	+ Nastavit společný program financování aktivit spojených s trvalou ochranou digitálních kulturních a vědeckých dat tak, aby bylo možné využít existující infrastruktury (CESNET, NK, NA) a zachovat možnost institucí rozhodovat o tom, jak budou jejich data ochraňována.
	+ Strategie by měla poskytnout rámec a obecná pravidla, neměla by obsahovat konkrétní technické požadavky ukazující na konkrétní řešení určitého problému (jako např. zajištění autenticity pomocí časových razítek a certifikátů) data, neměla by např. direktivně předepisovat formáty digitálních dat – rozhodnutí jaké formáty a signifikantní vlastnosti by měla nechat na expertech spravujících data (a globální komunitě) a producentech dat – jednotlivých institucích.
	+ Závazkem pro získání financí by měl být postup k certifikaci, resp. snaha vyhovět požadavkům DSA, stejně jako snaha o zajištění kvality procesů v organizaci (ISO 9001).
	+ To mj. vyžaduje specifikaci procesu a provádění certifikace digitálních repozitářů v institucích v gesci obou ministerstev.

## 4.2 Státem garantovaná datová centra pro zajištění bitové ochrany

* Vytvořit podmínky pro to, aby všechny instituce disponující velkými objemy dat, mohly získat nějaké pokročilé řešení pro základní ochranu bit streamu formou služby – tj. státem garantované datové centrum (nebo více datových center).
* IaaS pro storage a HW (může to být skrze spolupráci např. s CESNETem, případně vytvoření nové nebo využití jiné instituce).
* MK, MV a MŠMT by se měla dohodnout, zda jako klíčové HW úložiště bude dále rozvíjeno DÚ CESNET a budou mu delegovány další funkce, finance a odpovědnosti nebo zda stát paralelně bude budovat nové datové centrum. Financován by měl být přenos dat do těchto datových center, s tím, že by neměla klesnout podpora lokálně provozovaných řešení paralelně (z důvodů zajištění redundance, technologické diverzity atd.)
* Datové centrum musí splňovat standardy ISO 27000 a další průmyslové standardy. Toto datové centrum by mělo sloužit jako sekundární lokalita pro kohokoli nebo primární úložiště pro menší instituce. Musí mít zajištěno trvalé financování.

## 4.3 Státem garantovaný systém pro zajištění logické ochrany

* CESNET nebo nově budované datové centrum by měly pracovat směrem k rozšíření nabídky svých služeb – měly by provozovat tzv. long-term preservation (LTP) systém a s jeho pomocí začít nabízet certifikované řešení pro logickou dlouhodobou ochranu digitálních dat jako službu (DPaaS) pro instituce v gesci MK/MV/MŠMT.
* Datové centrum nebo datová centra by měla prostřednictvím LTP systému poskytovat služby dlouhodobé ochrany vyšší úrovně, logické. Financován by měl být přenos dat do těchto datových center, s tím, že by neměla klesnout podpora lokálně provozovaných řešení paralelně (redundance, technologická odlišnost).
* LTP systém poskytovaný jako služba by měl splňovat standardy ISO 16363 a ISO 14721 prokázané externí zveřejněným interním auditem a externí mezinárodní certifikací. Služby by měly používat především instituce, které nechtějí samy LTP systémy budovat, především menší instituce.
* Aby takové řešení mohlo existovat jako dlouhodobé důvěryhodné úložiště, potřebuje především garantované dlouhodobé financování, které nezruší kterákoli další vláda. Bez zajištěného financování nelze očekávat, že řešení projde externí certifikací podle ISO 16363.

## 4.4 Postupné budování sítě certifikovaných úložišť v jednotlivých institucích

* Větší knihovní instituce (MZK, Masarykova a/nebo Karlova univerzita, krajské knihovny a/nebo jejich konzorcia) by měly mít možnost získat financování na pořízení vlastního řešení pro dlouhodobou logickou ochranu (LTP systém).
* MK a MŠMT by měly ve svých finančních mechanismech vyčlenit finance na pořízení, zavedení a certifikaci LTP systémů. To bude vyžadovat také další standardizaci procesů v jednotlivých institucích, převody dat, a čištění dat a další náklady.
* Financování musí být zajištěno s oporou ve státním rozpočtu a rozpočtových kapitolách jednotlivých zřizovatelů (ministerstev) na delší dobu dopředu.
* Jedním z řešení financovaných z tohoto zdroje by mělo být rozvíjení řešení NDK na úroveň plnohodnotného LTP systému, který dokáže poskytovat logickou dlouhodobou ochranu jakémukoliv formátu dat, který je do systému vložen.
* Instituce by neměly být omezeny a měly by mít možnost si vybrat mezi svým lokálním řešením (LTP systémem) pokud mají zdroje na jeho vybudování, správu, provoz a rozvoj; anebo službou zmíněného datového centra vybudovaného pro účely bitové ochrany v části 4.2. V datovém centru by mělo být možné i např. uložit pouze zálohovou kopii dat.

## 4.5 Podpora Open Source řešení pro zajištění logické ochrany

* Adaptovat dostupný Open source LTP systém pro potřeby českých kulturních institucí a vytvořit “národní” verzi (distribuci) tohoto systému s doprovodnými aplikacemi.
* Na dalším vývoji by měla pracovat skupina IT odborníků z institucí, které systém implementovaly.
* K systému by měly existovat návody, procesní dokumentace, projektová dokumentace a také placená podpora od komerčních firem.
* Malé nebo střední instituce by mohly systém používat pro své digitální sbírky určené k archivaci.
* Mělo by se jednat o řešení založené na ověřených open source LTP systémech, které jsou úspěšně nasazovány v zahraničí, např. Archivematica, případně RODA[[16]](#footnote-16) - výhodou je mezinárodní komunita, neustálý rozvoj, komunitní podpora apod.
* Probíhající projekt LTP Pilot CESNETu by měl dát základ pro budování znalostní základny kolem systém Archivematica. Pokud se ukáže jako použitelný, mělo by být dále podporováno jeho používání v institucích.

## 4.6 Zřízení meziresortního metodického centra pro dlouhodobou ochranu

Metodické centrum by mělo:

* být garantováno vládou ČR a konkrétními ministerstvy (MK, MV, MŠMT);
* mít stálý rozpočet a být zřízeno jako individuální instituce nebo při konkrétní organizaci (NK, MZK, KNAV apod.);
* být obsazeno nezávislými odbornými pracovníky nezávislými na případné hostitelské instituci (tj. nepodléhají jejímu vedení, jsou vedeni, financováni přímo z rozpočtu ministerstev);
* vytvářet metodiky, národní strategie, národní standardy;
* vést vývoj “národních systémů nebo řešení”;
* podporovat kulturní instituce provádějící nebo začínající s dlouhodobou ochranou digitálních dat;
* spolupracovat se zahraniční LTP komunitou a jako svébytná entita se účastnit např. evropských projektů;
* zaštiťovat a podporovat proces certifikace digitálních repozitářů, překlady relevantních norem a metodik;
* být centrem plánování financování LTP aktivit na národní úrovni – vypisování a správa grantů aj.;
* vytvářet publikační činnost, provádět reportování vládě ČR apod.;
* spolupracovat s univerzitami v relevantních učebních sylabech.

## 4.7 Krajské knihovny jako producenti a správci digitálních dat

* Krajské knihovny musejí mít jasné plány a možnosti na uložení dat a jejich logickou ochranu, a to např. možnost využití určených datových center nebo využití open source LTP systému nebo DPaaS (viz výše).
* K dispozici jim bude metodika, podpora, strategie z národního metodického centra pro LTP.
* Musí mít možnost získání financování pro procesy dlouhodobé ochrany.

4.8 Malé instituce jako producenti dat

* Malé instituce, jako např. malé knihovny využívající VISK 7, musejí mít možnost využití určených datových center a využití open source LTP systému nebo DPaaS (viz výše).
* K dispozici jim bude metodika, podpora, strategie z národního metodického centra pro LTP.
* Musí mít možnost získání financování pro procesy dlouhodobé ochrany.
* VISK by měl poskytovat popis standardů LTP – doporučení vycházející z OAIS a dalších metodik, s jasně popsanými možnostmi řešení zajištění logické ochrany dat. Instituce by měly mít možnost využít finance VISKu na pořízení služby, uložení dat v jiné instituci, budování vlastního lokálního řešení, pořízení komerčního řešení.

## 4.9 Finanční mechanismy

* Financování ochrany digitálního kulturního dědictví jako samostatný program vlády ČR společně nebo pro jednotlivá relevantní ministerstva (MK, MV, MŠMT).
* Instituce musejí mít možnost požádat o několikaleté financování s následnou možností pokračování podpory na projekty uložení dat a jejich logické ochrany (ne vytváření dat). Nemělo by se jednat o financování nových pracovních pozic, prostředky musejí být získány reorganizací a evaluací stávajících procesů a aktivit v organizaci. Může se např. jednat o projekty na:
* instalaci LTP systému,
* pořízení HW infrastruktury v odůvodněných případech, kdy není možné HW sdílet nebo využít IaaS nebo DPaaS,
* převody dat do nových systémů, na nová média, ochranné akce na velkém množství dat aj.
* VISK by měl obsahovat podprogram týkající se dlouhodobé archivace. Měl by podpořit zavedení LTP systému (Open Source nebo nákup komerčního řešení), využití služby LTP a/nebo spojené náklady (příprava dat, převod dat, kontroly dat, realizace přenosu dat atd.).
* Financování digitalizace ve VISK by mělo být podmíněno prokázáním schopnosti zajistit dlouhodobou archivaci vytvořených dat. Každý žadatel by měl předložit plán archivace, kde přesně specifikuje, jak budou data dlouhodobě uchovávána. Bez schopnosti zajistit dostatečně bitovou ochranu výsledků digitalizace by nemělo být financování přiděleno. Instrukce by měla stanovit minimální požadavky na ochranu dat – umístění ve dvou lokalitách na dvou typech médií dostupných online, plus jedna offline záloha. VISK by měl tuto ochranu financovat, stejně jako financuje pořízení dat. Nelze spoléhat na to, že data budou ochraňovat instituce jako je např. NK, která v současnosti nedokáže externě prokázat kvality této činnosti a představuje pro zajištění ochrany digitalizovaných dat z VISKu podstatné riziko. Instituce by měly mít možnost se rozhodnout data do NK nedat vůbec, pokud prokážou schopnost je uložit v systému pro LTP jinde.

## 4.10 Organizační a systémové změny

* Knihovní zákon a následně mandáty jednotlivých organizací musí výslovně obsahovat a vyžadovat povinnost a zásady logické dlouhodobé ochrany dat. Je třeba doplnit knihovní zákon o jasnou odpovědnost knihoven za trvalou ochranu digitálních dat. V knihovním zákoně chybí specifikace typů dokumentů (digitální, WA, elektronické) a mělo by z něj být zcela jasné, že knihovny mají zákonem danou odpovědnosti za ochranu digitálního fondu. Zákon by měl být doplněn doporučením MK ČR, prováděcím předpisem, pro oblast LTP, kde budou zmíněny explicitně relevantní ISO normy a další obvyklé standardy pro oblast dlouhodobé archivace.
* Stávající zákony pokrývající tzv. povinný výtisk musejí být upraveny tak, aby bez jakýchkoliv pochyb zahrnovaly i born-digital dokumenty s povinností odevzdávání konkrétním knihovnám přes konkrétní mechanismy.
* Národní soustava povolání by měla být obohacena o pracovní pozice spojené se správou digitálních dat v knihovnách a dalších paměťových institucích. Především o kurátory digitálních dat, digitální knihovníky, správce digitální knihovny, metadatové specialisty a analytiky pro dlouhodobou archivaci.

# 5 Doporučení v oblasti zajištění kvality, standardizace a certifikace

## 5.1 Zajištění kvality; audit a certifikace

* Každá instituce provádějící dlouhodobou ochranu nebo která tento proces outsourcuje externí organizaci, musí pravidelně procházet auditem. Úspěšný audit by měl být podmínkou pro získání (nebo pokračování) financování. Výstupem auditu může být certifikace dle relevantních norem.
* Procesy v organizacích, které se zabývají dlouhodobou ochranou, musejí být zdokumentované a musejí odpovídat standardům. Ideální by bylo, kdyby meziresortní metodické LTP centrum formulovalo požadavky na zajištění kvality a zajistilo externí audit. Také dodržování procesů, standardů, metodik by mělo být podmínkou pro získání financování. Dokumentace a popisy procesů musejí být veřejně přístupné a otevřené, stejně jako např. technické řešení nebo smlouvy o outsourcování služeb.
* Základním orientačním bodem pro metodiku interního i externího auditu musejí být ISO normy (ISO 16363, ISO 14721) s tím, že realizace tzv. selfauditu nebo peer reviewied self-auditu by měla probíhat podle národní metodiky stanovené metodickým centrem pro LTP. V organizacích je také třeba také pracovat implementaci požadavků dalších souvisejících norem (ISO 27001, ISO 9001).
* Autoritou provádějící audity by mělo být metodické centrum pro LTP, které by disponovalo organizační platformou a možností najímat kvalifikované auditory.
* Získání oficiální certifikace podle ISO 16363 od akreditované certifikační autority nepovažujeme u všech projektů za nezbytné ani potřebné. Snad jen u velkých projektů financovaných z Evropské unie, případně národních projektů. Oficiální externí certifikace od certifikační autority bude poměrně finančně[[17]](#footnote-17) i časově náročná. Pravidelný self audit nebo externí audit podle na národní úrovni dohodnutých pravidel, spolu s publikací veškeré dokumentace k self auditu a hodnocení auditorů, bude pro většinu projektů dostačující zárukou snahy o dosažení statutu důvěryhodného dlouhodobého repozitáře ve smyslu ISO 16363.
* Při zavádění a provozování LTP systému je třeba postupovat podle společného evropského rámce pro certifikaci, který doporučuje postup ve třech krocích – selfaudit podle DSA, selfaudit podle ISO 16363 a externí audit podle ISO 16363. Nástroje jako DRAMBORA[[18]](#footnote-18) nebo PLATTER[[19]](#footnote-19) by měly sloužit při návrhu systému a analýze rizik, nebo pro metodickou orientaci v oblasti LTP. Je možné použít i německé ekvivalenty ISO 16363 a DSA (Nestor seal a Nestor criteria catalogue – DIN 31644:2012), nicméně národní metodické centrum nemusí garantovat stejnou míru podpory (standardy pro dokumentaci a doporučení, asistenci při selfauditu atd.) pro obě skupiny norem.

## 5.2 Doporučené postupy

Mezirezortní metodické centrum pro LTP by mělo také vypracovat doporučení v následujících oblastech:

* doporučení pro distribuci kopií dat mezi technologie a lokalitami,
* doporučení pro dokumentaci (změnový management, procesní management, configuration management),
* doporučení pro zajištění kvality řízení a organizace práce,
* doporučení pro standardy SW a worfklow pro zpracování dat – vycházející z normy OAIS.

# 6 Postup realizace koncepce

**1. Fáze**

* + společná strategie MK ČR, MŠMT, a MV
	+ vytvoření meziresortního metodického centra pro dlouhodobou ochranu digitálních dat
	+ změny knihovního zákona a prováděcí předpis
	+ analýza a pilotní využití Open Source systémů
	+ úprava knihovního zákona a zákonů týkajících se povinného výtisku
	+ obohacení národní soustavy povolání o relevantní pozice

**2. Fáze**

* + vybudování státem garantovaného centra pro bitovou ochranu kulturních dat
	+ vytvoření metodik pro LTP (bod 5.2. a 5.1.)
	+ vybudování systému pro výběr elektronického povinného výtisku
	+ úpravy finančních mechanismů (VISK, a další) tak, aby zahrnovaly problematiku archivace digitálních dat (bez plánu archivace nelze digitalizovat)
	+ vytvoření nových finančních mechanismů pro pořízení a budování LTP systémů
	+ vytvoření doporučení pro zajištění kvality v oblasti LTP (Normy, standardy, certifikace)
	+ asistence metodického centra při certifikaci
	+ prosazení standardů kvality archivace webu
	+ tlak na instituce, aby se věnovaly zajištění kvality dle ISO 9000 a ISO 27000

**3. Fáze**

* + vybudování státem garantovaného systému pro logickou ochranu dat, který bude schopno nabízet logickou ochranu dat jako službu
	+ financování budování lokálních řešení nebo využití služeb pro logickou ochranu digitálních dat
	+ financování podpory využití Open Source systémů pro logickou ochranu
	+ financování certifikace LTP činností

# 7. Současné aktivity

Ačkoli zájem o dlouhodobou archivaci se v ČR datuje již od počátku milénia (viz publikace [MACEK a WANNER, 2003]), do praktického budování LTP systémů se pustily zatím jen velké národní instituce ve velkých projektech (Národní digitální archiv, Národní digitální knihovna).

Některé aktivity vyplývající z předložené koncepce však již započaly a zcela jistě rozšíří dostupnost znalostí o dlouhodobé ochraně digitálních dat a podpoří snahu dalších institucí se touto oblastí seriozně zabývat.

Jmenovat můžeme např. projekt MZK, Masarykovy univerzity a sdružení CESNET s názvem LTP Pilot. Jeho cílem je otestovat LTP systém Archivematica, dostupný pod otevřenou licencí. V projektu bude posouzena funkčnost Archivematiky a budou prozkoumány možností integrace s prostředím digitálních úložišť CESNET. V rámci projektu bude instalováno několik instancí systémů, bude posouzena kvalita kódu, vytvořeny testovací scénáře vycházející z norem ISO 16363:2012 a ISO 14721:2012, posouzeny možnosti integrace se systémy, které se už v ČR používají; a systém bude testován s využitím reálných dat zúčastněných institucí. Protože zájem o problematiku dlouhodobé archivace mají i mnohé další instituce, (především Národní filmový archiv, Národní archiv, Knihovna Akademie věd, knihovna Univerzity Karlovy, a některé kraje) jsou součástí projektu také aktivity směřující s výměně zkušeností a informací (semináře, informační portál). Projekt začne v září 2014 a je plánován na 1,5 roku.

V oblasti certifikace digitálních repozitářů je aktivní MZK, která v současné době provádí pro Univerzitní knihovnu v Bratislavě posouzení jejich repozitáře pro externí audit dle ISO 16363.

Také paralelně probíhající aktivity v oblasti převodu norem (v knihovně UK vznikl překlad nástroje pro audit repozitářů *Data seal of approval[[20]](#footnote-20)*, v Národní knihovně vznikly překlady norem ISO 16363:2012 a ISO 14721:2012) nebo aktivity dalších projektů (Projekt NAKI Knihovny Akademie věd Česká digitální knihovna, projekt NAKI Národní knihovny týkající se e-depositu) bezpochyby přispějí dosažení cílů vytyčených v této koncepci.

# 8. Použité zdroje

**FOX, Lisa L. 1985**. Archival Preservation in the Age of Technology *Provenance, Journal of the Society of Georgia Archivists*: Vol. 3, Iss. 1, Article 4. Dostupné z: <http://digitalcommons.kennesaw.edu/provenance/vol3/iss1/4/>.

**GARRETT, John a Donald WATERS. 1996.** *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* [online]*.* The Commission on Preservation and Access and The Research Libraries Group, 1996 [cit. 2014-05-17]. 64 s. Dostupné z: <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>.

**MACEK, Oskar a Michal WANNER. 2003.** Strategie digitálního archivování. *Archivní časopis*, č. 53, 2003. Zvláštní příloha. 64 s. ISBN 80-86466-03-5.

1. <https://ec.europa.eu/digital-agenda/en/digitisation-digital-preservation> [↑](#footnote-ref-1)
2. Knihovním dokumentem [je] informační pramen evidovaný jako samostatná jednotka knihovního

fondu knihovny. Knihovním fondem [je] organizovaný, soustavně doplňovaný, zpracovávaný, ochraňovaný a

uchovávaný soubor knihovních dokumentů, [↑](#footnote-ref-2)
3. podporovanou z veřejných rozpočtů od konce 90. let - EU projekty počínaje 5. rámcovým programem - FP5 (1998-2001; projekty ERPANET, DigiCULT, PRESTO, ECHO, MINERVA, FP6 (2002-2006; PRESTOSPACE, DELOS, BRICKS, CASPAR, DPE, PLANETS, PRESTOSPACE) a FP7 (2007-2013; PROTAGE, DL.ORG, LIWA, SHAMAN, KEEP, PRESTOPRIME, SCAPE, ARCOMEM, TIMBUS); americké projekty jako NDIIPP nebo novozélandský NDHA a další národní projekty [↑](#footnote-ref-3)
4. Dočasně ukládaná data musí být možné nahradit (přeskenováním, konverzí z masterů atd.). [↑](#footnote-ref-4)
5. Tedy konkrétně mít procesní dokumentaci; strategii; plány pro různé situace; dokumentaci pravidelně revidovat apod. [↑](#footnote-ref-5)
6. <http://ndk.cz/narodni-dk> [↑](#footnote-ref-6)
7. <http://www.nkp.cz/soubory/ostatni/tz-ndk-google.pdf/at_download/file> [↑](#footnote-ref-7)
8. <http://www.manuscriptorium.com/> [↑](#footnote-ref-8)
9. Metadata již z procesu digitalizace obsahují základní technická a administrativní metadata, záznamy o procesech digitalizace apod. Tyto údaje pak mohou posloužit pro procesy dlouhodobé ochrany nebo jako vstup do LTP systému. [↑](#footnote-ref-9)
10. http://visk.nkp.cz/2013/VISK7-podm2013.doc [↑](#footnote-ref-10)
11. <http://www.enumerate.eu/fileadmin/ENUMERATE/meetings/ENUMERATE-FocusGroup-on-Digital-Preservation-Monitoring_TheHague_20130117_v01.pdf>) [↑](#footnote-ref-11)
12. <http://www.digcur-education.org/> [↑](#footnote-ref-12)
13. <http://www.datasealofapproval.org/en/> [↑](#footnote-ref-13)
14. <http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html> [↑](#footnote-ref-14)
15. Jako je např. LTP systém Preservica ve své cloudové verzi, http://preservica.com/editions-pricing/. [↑](#footnote-ref-15)
16. <https://www.archivematica.org/wiki/Main_Page> a <http://www.roda-community.org/> [↑](#footnote-ref-16)
17. Proces certifikace dle ISO 16363 podle prováděcí normy ISO 16919 bude vyžadovat auditorskou firmu, která bude akreditována u certifikační autority tj. Českého institutu pro akreditaci, a.s. (<http://www.cia.cz/>). Tato firma bude dle ISO 16919 mít nejen akreditaci k ISO 16363, ale také k ISO 17021 Posuzování shody - Požadavky na orgány provádějící audit a certifikaci systémů managementu. [↑](#footnote-ref-17)
18. http://www.repositoryaudit.eu/ [↑](#footnote-ref-18)
19. http://www.ndk.cz/platter-cz [↑](#footnote-ref-19)
20. <http://www.datasealofapproval.org/en/> [↑](#footnote-ref-20)